**Data Bases: Flat File Vs Relational**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Read the following problem:

Mary Carey runs a small CD rental firm. Initially she had a card system with a card for each CD, containing the details shown on the sample card below. If the firm had more than one copy of a CD, a separate card was created for each.

When a video was borrowed, the borrower’s name phone number and the date was entered onto the card. Once it was returned the name was crossed out. The cards were stored in order of CD title; the CDs on the shelves were arranged in the same order.

|  |  |  |
| --- | --- | --- |
| ***Hullabaloo CD Hire*** |  |  |
| CD Title | Let’s ride |  |
| Performer | Front End Loader |  |
| Category | Rock |  |
| Released by | Shagpile |  |
| Date Purchased | 12/7/07 |  |
| Cost | $19.95 |  |
| Borrowed by | Phone Num | Date |
| *~~G Smith~~* | *~~9321 4567~~* | *~~13/8/07~~* |
| *~~P Jones~~* | *~~9856 6789~~* | *~~15/10/07~~* |
| *~~J Jenson~~* | *~~9345 3434~~* | *~~11/11/07~~* |
| *~~G Smith~~* | *~~9321 4567~~* | *~~23/2/08~~* |
| *M Monty* | *9111 2222* | *5/5/08* |
|  |  |  |

Mary was keen to provide a fast and efficient service so an assistant was hired when the business became more popular. Despite this, she found that customers were kept waiting and that requested CDs were often difficult to locate.

**Flat File Database**

The above file is an example of a flat file database where all the data is contained in one file or table. A flat file must have no repeating groups, that is, no field can have more than one value in a particular record.

**Questions:**

The problems with the above card or file system are that it is hard to find the card of the movie in a quick time. The CDs were also hard to locate. By crossing out the people who have returned the movie this makes the card very messy and hard to read. The details of the people who borrowed the movies cannot be kept properly because it is crossed off every time a movie is returned. Also it is hard to locate who hasn’t returned over due movies.

If the persons details change than there will be no record of this until they next borrow a movie because the phone number is crossed of everytime a movie is returned.

In the above example there is repeating fields because different customers may have the same last or first names so this could be repeated.

**Relational Database**

A relational database is a database consisting of more than one table/file linked together with the use of primary keys or common fields.

Primary keys are fields with a unique number which can identify a record.

The advantages of a relational database are:

* reduced data redundancy – unnecessary duplication of data
* improved data integrity or accuracy of data
* able to access the data from separate but related tables or files very quickly and simply
* allows easy querying, reporting, etc.CDs were not returned promptly. Often the borrower gave an incorrect phone number.

To help overcome this problem, Mary introduced a **second card file of customers’ details.** Each card contained the surname, given name, address, phone number and whether the customer was under 18. Cards were stored in alphabetical order by surname.

Mary decided to start by transferring the details from the **customer** card file to a database file.

**Questions:**

The advantage of including a field, customer name, which contains a unique integer are that it will be easy to find the customer details if they are overdue on returning a movie, if they are borrowing or wanting to change their details. Also by having weather the customer is over 18 or not will show weather they are trustworthy at bringing back the movie.

The appropriate data types for each field are:

Firstname text

Surname text

Address text

Phone number numerical

Customer 18 or not text

Film most recently borrowed text

Number of films borrowed numerical

She finds there is a problem transferring the data from the **CD** card file. The problem is related to the structure of the data on the card which is inconsistent with that of a “flat-file”.

The problem is that the structure of the CD is not consistant and might contain some wrong data information which will make the data incorrect. Because flat files contain rows and columns that has no repeating groups whereas some groups will need to be repeated like surnames and names.

Mary cannot decide how to set up the **loan** file. She realises that to be able to quickly identify and contact customers with overdue CDs she might need to store details such as name, address, phone in the loans file as well as the customers file. Her friend advises that she should avoid data redundancy where possible.

1. Data redundancy means that building in extra items that are not strictly necessary for something to function.
2. Data redundancy should be avoided so that there is not too much information on the computer and this will save room for other things. By not using too many extra fields this will save confusion on the data base.

The fields that Mary must include in the loans file is:

* CD Title
* Preformer
* Cateogory
* Released by
* Date purchased
* Cost
* Borrowed by
* Number of copies in store

Mary finds that there are still a number of problems. To print a list of the titles of the most popular CDs, requires her to access the loans file to get the CD number and then the CDs file to get the title and artist. She has heard that a relational database would allow her to do such a task much more simply.

A relational database will help solve Marys problem because a layout could be made that said the most popular Cd’s on it. This can be accessed from the home page, where there would be a link.

A relational database is to consist of three tables:

* Athletes, contains details about each athlete, (ID, firstname, surname, country)
* Events, contains details about each event, (ID, eventname, sport, venue)
* Program, contains details about who is in each event and when it is on, (athlete ID, eventide, date, time)

Use the diagram below to show the relationship between the three tables.